

FUNCTIONAL DIFFERENCES IN THE STORAGE OF WORDS
AND NONWORDS IN LONG-TERM MEMORY

By

GARY BRUCE FORBACH
//

Bachelor of Arts
University of Denver
Denver, Colorado
1969

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1972

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1973

Thesis
1973D
F692f
cop. 2

FEB 15 1974

FUNCTIONAL DIFFERENCES IN THE STORAGE OF WORDS
AND NONWORDS IN LONG-TERM MEMORY

Thesis Approved:

Robert F. Stammers
Thesis Adviser

John Battle

Robert J. Usher

Larry Hochhaus

D. D. Burken

Dean of the Graduate College

ACKNOWLEDGEMENTS

The bumpy course of the development and implementation of this thesis has been considerably softened by the guidance and counsel of a number of people. A great debt of gratitude is owed Dr. Robert F. Stanners, who served as thesis adviser, for both his constant assistance and the use of his research facilities. Drs. Robert J. Weber and Larry Hochhaus have quite often made themselves available for discussion of solutions to nagging problems, and have often provided moral support when it was needed. To Dr. John H. Battle, a special thanks for his help in consideration of the problem from the linguist's point of view. Finally, I would like to thank my wife, Paula, for her much appreciated typing skills, her patience, and her continued support of my academic endeavors.

This thesis is dedicated to the hope of an improving world, the advancement of human understanding, and the development of the science of Psychology.

TABLE OF CONTENTS

Chapter	Page
I. STATEMENT OF THE PROBLEM	1
Purpose of the Study	1
Relevant Literature	2
Frequency Variables	2
Homography	6
Encoding Variables and Homography	9
Decision Aspects of Classification	12
Semantic Variables	13
Content-Addressable Models	20
A Random Retrieval Model	23
Immediate Antecedents of the Study	24
Rationale	25
II. METHODOLOGY	27
Subjects	27
Apparatus	27
Materials	29
Procedure	30
Selection	30
Learning	31
Test	33
Experimental Design	33
Scoring of Data	33
III. RESULTS	35
Selection Session	35
Learning and Test Sessions	38
IV. DISCUSSION	43
Selection Session	43
Learning and Test Sessions	45
V. SUMMARY	49
A SELECTED BIBLIOGRAPHY	51
APPENDIX A - MEANINGFULNESS RATING INSTRUCTIONS FOR R WORDS SELECTION	54

Chapter	Page
APPENDIX B - STIMULUS ITEMS USED IN EACH CLASSIFICATION SESSION	55
APPENDIX C - INSTRUCTIONS TO SUBJECTS FOR SELECTION SESSION	57
APPENDIX D - INSTRUCTIONS TO SUBJECTS FOR THE LEARNING TRIALS	59
APPENDIX E - SAMPLE LAYOUT OF CARDS USED FOR LEARNING, AND STUDY DEFINITIONS	61
APPENDIX F - INSTRUCTIONS TO SUBJECTS IN LEXICAL DECISION STUDIES	63

LIST OF TABLES

Table		Page
I.	Mean Latency, Mean Error Latency, and Number of Errors in Each Category for the Selection Session Classification Trials	36
II.	Summary of Significant Differences in Category Means, as Tested by the Newman-Keuls Test	37
III.	Mean Latency, Mean Error Latency, and Number of Errors in Each Category for the Learning and Test Session Classification Trials	39

LIST OF FIGURES

Figure		Page
1.	Mean Latency of Response for the Experimental and Control Groups Immediately After Learning R Words (Learning), and Then 48 Hours Later (Test)	40

CHAPTER I

STATEMENT OF THE PROBLEM

Purpose of the Study

The present study is an attempt to determine some of the processes involved in word recognition. It will be primarily concerned with the manner in which words are stored in memory. In particular, the present study is concerned with possible differences in the method of storage of words and nonwords (nonsense words) in relatively permanent memory. For example, these differences might be in terms of the location in which the items are stored. (In this text, location refers to a hypothetical subdivision of memory, and not a specific physical locale.) The nature of the production and understanding of words suggests that an orderly, interrelated arrangement of memory exists for words. Are nonsense words introduced into this network if Ss are required to learn them? Or are they stored apart in a special type of memory storage? Data processing machines create temporary storage registers or "scratch files" as they are required. Analogous to this may be the temporary learning of a name, address, or telephone number needed for a day or two, but forgotten after it is no longer useful. The intent of this discussion is not to suggest that because some material is remembered and some is forgotten that there must be permanent and temporary memory storage locations, but rather that such an organization could be functional. If memory for nonwords is functionally different from memory

for words, the results of studies using nonwords to determine memory characteristics should not be used to infer characteristics of memory for words. The present paper is concerned with examination of evidence which supports or does not support the hypothesis of different storage locations for different types of material to be stored. In addition, the present study is expected to replicate some patterns of results of previous studies and lend support to earlier hypotheses about the characteristics of word memory. A review of some relevant research will define these expectations, and lead to the rationale for the present study.

Relevant Literature

Frequency Variables

A paper by Stanners, Forbach, and Headley (1971) examined the effect of frequency of initial and terminal letter of trigrams on word/nonword classification latency. Three categories of material were used: consonant-vowel-consonant words (WORDS), consonant-vowel-consonant nonwords (CVCs), and consonant-consonant-consonant nonwords (CCCs). Identical initial and terminal consonants were used in the three categories of material, e.g., SAT, SUT, and SBT. By varying the frequency of the initial and/or terminal consonant (high or low), four sets of items were formed. They found that the CVCs take the longest to classify, CCCs take the least time, and WORDs take an intermediate amount of time; but each category was different from the other two. Letter frequency did not produce a significant effect on CCCs. However, letter frequency (component frequency) did produce a significant effect on CVCs and WORDs. The results were interpreted as supporting a two stage

process. Since there was no effect of frequency on CCCs, it was assumed that the illegality of the CCC produced the rejection of the item at some early encoding stage. Further, it was assumed that if the item is lawful, it is encoded and a search through memory takes place. This search was assumed to proceed through a subset of the internal lexicon, where the size of the subset is a function of the frequency of the components, in this case, the initial and terminal consonants. If the item is a CVC, unsuccessful exhaustive search of the subset is assumed to occur. Thus for a CVC, search continues longer than for a WORD, which does not require exhaustive search, since it can be classified as a word when a "semantic marker" is encountered. At this point, Stanners et al. (1971) noted that an important question was whether the latency differences between CVCs and WORDs should be attributed to individual letter frequencies or the frequency of the encoded version of the whole item. A search through the Mayzner and Tresselt (1965) count indicated that frequency of the initial and terminal letters covaried with frequency of the CVCs as units in English. (It should be noted that CVCs can occur as words, e.g., "cat," and as units which occur as parts of other words, e.g., "category," "concatenate." Also, CVCs may occur only as parts of other words, e.g., "swamp," "category," and "concatenate.") Thus although the data suggested some sort of search of a subset of memory for an encoded form of a WORD or CVC, it could not differentiate between the following possible alternatives:

(1) The subset is delimited by the frequency of the components of the item, i.e., the initial and terminal letters; or (2) the subset is identified by the encoded version of the item as a unit. However, whatever the method of specifying the subset for search, exhaustive

search of a subset of memory for CVCs, versus partial search through a subset of memory for WORDs can account for the differences in latency.

To choose between the unit-frequency hypothesis and the letter-frequency hypothesis, Stanners and Forbach (1973) did a study using as materials five-letter items. There were three categories: consonant-consonant-vowel-consonant-consonant words (WORDs), consonant-consonant-vowel-consonant-consonant nonwords (CCVCCs), and consonant-consonant-consonant-consonant-consonant nonwords (CCCCCs). Across categories, identical initial and terminal consonant clusters (CCs) were used. These CCs were varied from high to low frequency. Since the frequency of the CCVCCs and CCCCCs as units (that is, the frequency with which the S had experienced the five letters in a given order) is the same for all the items, essentially zero, then any differences in classification latency within each category must be attributable to the frequency manipulations of the initial and terminal CCs. For English words, the problem is complicated, since in addition to the frequency of the CCs, there is the frequency with which the S has encountered a given word. This frequency is assumed to be approximated by the Thorndike-Lorge (1944) count. Thus within the category WORDs, frequency of CCs alone may not account for differences in latency of classification.

The results showed a significant difference in classification latency between all categories, with CCCCCs requiring the least time, CCVCCs requiring the most time, and WORDs requiring an intermediate amount of time. More interesting was the significant effect of CC frequency on latency which occurred in all three categories of material. Generally, for the nonwords, low frequency CCs produced faster classification latency, and high frequency CCs produced longer latencies. The

effect for WORDs was generally opposite that for nonwords. But since WORDs have a unit frequency in addition to CC frequency, results for them might be expected to differ from those for nonwords.

These results were interpreted as follows: (1) The information available from the CCs is abstracted and used to functionally circumscribe a subset of memory for search. (Rubenstein, Lewis, and Rubenstein, 1971a, have proposed a similar process which they have called "marking.") (2) After marking is completed, the item is encoded as a unit for comparison with the contents of the marked subset of memory. If the item is unlawful as in the case of the CCCCQs, it is rejected immediately. If not, it is encoded and search of the marked subset is carried out, exhaustively for CCVCCs since they are not in memory. On the other hand, for a word, there is only partial search of the subset. This process would account for the differences in latencies for the three categories of material. (3) It is also assumed that both marking and search time vary directly with the frequency of the CCs, since the high frequency CCs represent more occurrences in English than do the low frequency CCs. This would account for the differences within the two categories of nonwords. (4) To account for the opposite effect of CC frequency for WORDs, examination of the unit frequency of the WORDs in conjunction with the CC frequency is necessary. A set of 13 filler items were used for the comparison. These items were constructed with identical high frequency initial and terminal CCs, differing only in the medial vowel and the Thorndike-Lorge frequency, either very high or very low. For example, CROSS-CRASS, and TRUTH-TROTH. The means of these sets showed high frequency items were classified significantly faster than low frequency items. This was interpreted as an indication

that search through a subset is ordered, with high-use items always searched first, and low-use items searched only if necessary.

A study by Loftus, Freedman, and Loftus (1970) using a different task (retrieval of an instance of one of many different semantic categories) has shown that high frequency words tend to be given first. This could be interpreted as evidence that given a general area of memory for search, the most frequent items would be attended to first. This interpretation is also congruent with the Spew hypothesis proposed by Underwood and Schulz (1960). They cite as suggestive evidence classical free association data (Johnson, 1956, and Howes, 1957), and category association data (Cohen, Bousfield, and Whitmarsh, 1957). These studies indicate that the frequency of a word is directly related to response probability.

Homography

An earlier paper by Rubenstein, Garfield, and Millikan (1970) had found some of the same results reported in Stanners et al. (1971) and Stanners and Forbach (1973), and some other pertinent results. The task in the Rubenstein et al. (1970) study was also word/nonword classification. They found that words were classified faster than nonwords, and high frequency words were classified faster than low frequency words. In addition, they found that homographs (words with multiple meanings) were classified faster than nonhomographs, and that within homographs concreteness had a significant effect. No satisfactory explanation of the concreteness effect was found. However, after they ruled out the possibility that lower reaction time (RT) for homographs could be attributed to greater familiarity, they postulated that a

possible explanation for the effect of homography could be the existence of multiple entries--that is, an entry for each different meaning--in the internal lexicon. This effect was assumed to exist for unsystematic homographs (homographs whose meanings are unrelated, such as "yard"--a linear measure, and "yard"--an enclosure). Thus if no particular meaning was referenced, as in a word/nonword classification task, words with multiple entries would have an advantage since the most accessible entry could be referenced. They also assumed that systematic homographs, such as "plow" (the verb) and "plow" (the noun) would probably not have multiple entries, but would instead be represented in the internal lexicon by one entry, and a special "label" for each syntactic class required. Post hoc partitioning of the appropriate homographs tended to support this interpretation. There was also an indication that relative frequency of meaning for homographs might be important for the facilitation effect of homography. That is, if one meaning of a homograph is much lower in frequency relative to the other, it was suggested that the effect of homography might be minimized.

To account for the observed effects, Rubenstein et al. (1970) postulated a search process which involves at least four stages:

- (1) Quantization, the division of the stimulus into segments;
- (2) marking, a process which uses the output of quantization to mark some subset of lexical entries;
- (3) comparison of quantization outputs with marked entries; and
- (4) selection of the marked entry that satisfies the accuracy criterion set by S.

Steps 1-3 are assumed to proceed as long as required. Also, marking of one output of quantization is assumed to proceed while subsequent quantization is continuing, as required. The results were explained as follows: (1) The word

frequency effect is the result of marking entries in the highest frequency range first, and then proceeding to lower frequencies as required; (2) the homographic effect is due to a random search within each marked set. The probability of finding one of several entries for a homograph is higher than for a nonhomograph with a single lexical entry. Thus RT should be faster for homographs on the average; (3) the suggested effect of relative frequency of homographs was tentatively attributed to the assumption that different frequencies are marked at different times in the search process, and thus for homographs with different relative frequencies only one entry is available during comparison; (4) nonwords require the most time since exhaustive search of all sets marked by quantization output is required.

A second study (Rubenstein et al., 1971a) examined the effects of systematicity and relative frequency of meanings of homographs in much more detail. They found that the facilitating effect of homography is observable only when the meanings of the homographs are not systematically related, and also tend to be equiprobable in relative frequency. These findings support the tentative model suggested in the earlier paper and suggest that marking proceeds in order of frequency of occurrence of meanings if there are multiple meanings of a word. Since systematic homographs had no significant facilitation effect, the hypothesis that they do not have multiple entries was supported. The authors also noted that the idea of a random search may be strange. But since word recognition usually occurs during conversation or reading, additional information is available which helps reduce the size of the set to be marked to a very few items. A random search of a small number of items in the search set would take very little time, and therefore the idea of random search may not be too unreasonable.

Encoding Variables and Homophony

An explanation of word storage and recognition must also consider the nature of the encoded form of the stimulus which is to be compared with the internal form of the stored word. Rubenstein, Lewis, and Rubenstein (1971b) attempted to ascertain whether or not the stimulus is recoded phonemically in visual word recognition, and if so, when in the recognition process the recoding is accomplished. They also attempted to determine whether search through the internal lexicon employs a phonemic or orthographic code.

Experiment I involved further analysis of the nonwords used in Rubenstein et al. (1971a). There were three categories of these nonwords: (1) Items which were orthographically illegal in English, but pronounceable (gratf, lamg); (2) items which were orthographically illegal and unpronounceable (likj, crepw); and (3) items which were orthographically legal (strig, plind). The authors proposed that if phonemic recoding occurs, there should be a difference in the two types of illegal items due to difference in pronunciability. They also proposed that if latency for legal nonwords is greater than that for both the illegal types, phonemic recoding can be attributed to quantization. This is based on the assumption that the illegal items would be rejected before any search would be required. The results showed that illegal unpronounceable nonwords took significantly less classification time than did the illegal pronounceable items. Legal nonwords took the longest. Both results support the prediction of the model.

Experiment II of the Rubenstein et al. (1971a) study was designed to determine the nature of the encoded form of the stimulus used for

search through the internal lexicon. Three types of nonwords were used: (1) Nonwords homophonic with low frequency words, e.g., stail; (2) nonwords homophonic with high frequency words, e.g., brane; and (3) nonwords not homophonic with words. The authors predicted that if marking and comparison of memory entries is carried out using phonemic representations of the stimulus, then the phonemic match of "brane"/breyn/ would match the phonemic match of "brain"/breyn/. Thus correct classification of "brane" would require the usual exhaustive search time, plus time to compare orthographic representations of an item whenever a phonemic match was found. However, nonwords not homographic with words would require only exhaustive search time, and should therefore on the average take less time to classify. The results showed a significant effect of homophony in the predicted direction. This result supports the phonemic recoding hypothesis also.

Experiment III examined effects of frequency and homophony for words. In this case, homophones should require more time for classification than nonhomophones, due to the delay caused by finding an inappropriate orthographic entry. Frequency should also have a significant effect, i.e., high frequency words should have faster latencies. The results supported the predictions, and were interpreted as an indication that a match in phonemic code is not sufficient to classify an item; orthographic information is checked also, and thus if the first match found is orthographically inappropriate, recognition is delayed. Rubenstein et al. (1971b) also suggest a possible method of detection of phonemic illegalities based on distinctive features analysis during quantization. In their experiment, the illegality occurred only as the final phoneme cluster. They found generally that as the number of illegal distinctive features increased, the illegality

tended to be discovered more rapidly. However, this interpretation was not able to explain why these illegal nonsense words were not classified faster than words. If the detection of illegalities occurs during quantization as they propose, illegal nonsense words should have shorter latencies than English words. Stanners and Forbach (1973) found that illegal nonsense words (BRKNG, PRBSS) were classified faster than any other material. This supports the hypothesis that detection of illegalities occurs during quantization, but suggests that quantization may be sensitive only to very gross illegalities in orthography, and that the definition of illegality used by Rubenstein et al. (1971b) may be the cause of the apparent inconsistency.

An experiment by Kollasch and Kausler (1972) using aural presentation of study words for later visual recognition, demonstrated an interesting effect of homophony. The miss rate for recognition of the low frequency, or secondary form of the homophone was greater than for the high frequency, or primary form of the homophone. Since the study trials were on the phonemic representation of the item (the same for both forms), recognition rates should not differ unless the "tag" assigned to an item in memory to mark it for later recognition was based on a priority system which varied with frequency. The results indicated that the most frequently used item was the one which was usually "tagged." The S hearing /payn/ was more likely to recognize the visual presentation of pain than paine. Although this study was not a word/nonword classification study, it supports the model proposed by both Stanners and Forbach (1973), and Rubenstein et al. (1971a, 1971b) with regard to predicted frequency effects. It does not, however, differentiate the two models with respect to the locus of the facilitation effect of frequency.

Decision Aspects of Classification

A problem which has not been considered in much detail involves the decision aspects of classification. The models discussed have assumed that some kind of exhaustive search of at least a portion of the internal lexicon is the basis for classification of a nonword correctly. However, what happens at the end of that fruitless search may be quite important to retrieval models in general. For example, when a search is finished, is a second search started immediately for a "recheck" of the results before a response is made? Does a bias toward faster responses change part of the process? Will the S who feels accuracy is most important do the task with more checks and rechecks? Does practice on the task affect words and nonwords the same way? Although this last question is usually answered in the affirmative (or apparently assumed so by most experimenters), there is evidence that practice affects correct word classification, and correct nonword classification differentially. Forbach, Stanners, and Hochhaus (1973) have shown that over the course of approximately 300 classification trials, word classification speed shows no significant improvement. On the other hand, practice on the classification task produces a large reduction in the latency of classification for nonwords. This effect appears to continue to some extent over even the last 100 trials. One interpretation of this effect is based on possible courses of action adopted by the S after an unsuccessful search. For example, he may decide to "recheck" the search subset in case he missed the item during the initial search. Practice on this task may lead to more confidence in accepting the outcome of the first search, and eliminate both the time for a second search and the time to decide whether or not another

search is required. This interpretation would predict no decrease in latency for words as a function of practice, since no decision regarding "rechecking" would be necessary on all those trials in which a word is found. The presence of an interaction of practice with stimulus type (word vs. nonword) for a lexical decision task suggests care is necessary to avoid confounding practice with other independent variables.

Semantic Variables

One question which has not yet been considered involves the extent to which semantic information is involved in the task of classifying an item as word or nonword. Does indexing a location in the internal lexicon automatically provide semantic information also, or is another step necessary to retrieve semantic information? Would semantic information about the to-be-found item aid its indexing? What other kinds of information are being assimilated by the S, and does this information facilitate or interfere with the task?

Stroop (1935) found that if subjects were instructed to read the color names "red," "green," "blue," "brown," and "purple," when those words were printed in ink of a color different from that named by the word, they could read the list as quickly as a control list printed in black ink. However, when asked to state the color of the ink in which each of a list of words was printed, there was a large interference effect. (Here also, the color named by the word was always different from the color of the ink in which it was printed.) Naming the color of the ink used to print each word on the list took much longer than naming the color of blocks of squares or swastikas which served as the

control list. It appears that even though he was not asked to interpret the actual letters (which formed words), the S was extracting semantic meaning from the letter strings. And in fact, although practice reduced the interference effect, S could not completely inhibit the normal response to a letter string and do the task as fast as the control task.

If such a strong effect as this can occur even when S tries to suppress it, it might be the case that there are other more subtle, but powerful effects on a subject's performance when instructions make no effort to control or manipulate these effects.

A study by Meyer and Ellis (1970) has suggested that in some tasks, the S might be engaged in parallel searches for some specific semantic information and/or evidence that a given letter string is in fact a word. In their experiment, the S was required to classify a letter string as (1) a word or nonword, or (2) a member (or not a member) of a small or large semantic category when given the category name. An S could classify a word correctly faster than a nonword correctly, as found by previous investigators mentioned earlier. In addition, however, he could classify a nonword as not a member of a semantic category faster than he could classify a word as not a member of a semantic category. To account for these results, Meyer and Ellis proposed a parallel race between a "meaning-decision," and a "word-decision." The time for each of these decisions is assumed to be a random variable, with the two distributions overlapping somewhat. (The proposed model does not specify the relationship between the means of the two distributions.) Thus for nonwords, when the task requires a decision about meaning, both the word-decision and the meaning-decision enter the race. The conclusion of either will produce the required "no" response, since if the letter

string is not a word, it cannot be in a semantic category. When a nonword is the stimulus for the question "Is this a word?" only the word-decision enters the race. Thus on the average, this search would take longer (only the word-decision enters the race, and the outcome must wait until it finishes). The same explanation applies for words, but since no exhaustive search is required, words would require less search time than the nonwords for a word-decision. On the other hand, when a word is presented for semantic search, but is not a member of the semantic category cued, the S must wait for the results of the "meaning-decision," even if the word decision finishes first. Thus negative semantic decisions take longer for words than for nonwords. These results suggest it might be useful to consider some parallel process models in attempting to explain word classification tasks.

Using a word/nonword classification task, Meyer and Schvaneveldt (1971) demonstrated a facilitation of the classification of a word which immediately followed classification of a semantically associated word, e.g., DOCTOR-NURSE. The S saw a display of two letter strings displayed one above the other. In the first experiment S was asked to identify both strings as words or not. When two associated words were displayed together, classification time was significantly faster than the response latency for classifying two unassociated words. However, the "no" responses were somewhat confounded since a nonword discovered in the first position of the display terminated evaluation and produced a fast "no" response. In Experiment II the same stimuli and apparatus were used, but S was asked to respond by pressing the "same" key if both letter strings were words or both were nonwords. He was to press the "different" key if a word and a nonword appeared in the display.

Again, associated words required less time for correct classification than did unassociated words. Also, the "same" response took longer for two nonwords than for two words of either type. The results were explained in terms of a model involving two separate, successive decisions. Serial processing from the top to the bottom of the display is assumed. Each letter string is evaluated as to being a word or nonword, and then the decisions are compared, producing the correct response if each lexical decision is correct. Since nonwords typically require more time for a single correct classification than words, two such responses should require a longer time than two classification responses for words. However, no previous work has provided a possible explanation for the facilitation effect of associated words. The authors suggested a "location-shifting" model which assumes semantically related items are functionally proximal, and the time required for a shift from one word to another varies directly with semantic relatedness. This explains the word effects. The authors also suggest a characteristic of nonword memory search which is an aspect of the model proposed by Stanners and Forbach (1973), namely, a nonword is classified correctly not by searching all of the internal lexicon, but rather by searching the memory subset in which a particular item would be stored if it were in fact a word. This requires exhaustive search of at least some area, which on the average would require more time than search for a word found before the termination of an exhaustive search. This aspect of the model accounts for the nonword "same" results.

One other interesting result reported by Meyer and Schvaneveldt (1971) is the fact that when the word was displayed on top of a mixed display ("different" response required), response latency was shorter

than when the nonword was displayed above a word. By invoking the often replicated word frequency effect, the "location-shifting" model explains these results as follows: Lexical-decision search starts first in the more frequent sectors of the internal lexicon; if an item is not found, a shift is made to another sector, and so on until the item is found, or search of possible locations is complete. Thus if a word is evaluated first, the "pointer" is closer to the beginning point (especially when the word is a frequent word), and little or no shift time is required to process the second item, the nonword. However, when a nonword is first, the pointer shifts from the high frequency sector, to the low frequency sector, and then back to the high frequency sector when the second item is processed. This requires more time, and thus when the nonword is displayed first, the total decision time is longer.

A subsequent paper by Schvaneveldt and Meyer (1971) using a display of three strings of letters has replicated the association facilitation effect, and produced some critical tests of possible models. The S was asked to decide if all three letter strings were words ("yes") or not ("no"). Three types of letter strings were used: (1) Words unassociated with any other words in the display; (2) words which were associated with another word in the display; and (3) nonwords. Displays with three nonwords or three associated words were not used. An important variable was the position of a nonword in a display with two words, and the position of a word in a display with two nonwords.

Three alternative models were considered, A Spreading-Excitation Model (Collins and Quillian, 1970; Warren, 1972) which is based on the concept of neural excitation. It is assumed that retrieval of an item excites surrounding locations, and thus facilitates retrieval of

information from these locations. The Location-Shifting Model (Meyer and Schvaneveldt, 1971) proposes that information can be "read out" of only one memory location at a given instant (much like information is read from a magnetic tape by a computer). They assumed that time is required to shift locations, and since associated information is assumed to be stored closer together on the tape, associated material would require less time because of less shifting. A third possible model is a version of a Semantic-Comparison Model. This model suggests that the association effect is due to lowering the response criterion for associated words during the early evaluation of an item (Schaeffer and Wallace, 1970). This bias toward the positive response is assumed to then facilitate a positive response for associated words, and inhibit a negative response for associated words. (However it has been argued that this type of model may not be appropriate for a lexical decision task, in Meyer and Schvaneveldt (1971) since in the lexical decision task only information about word/nonword status is required. If another stage for meaning comparison is added, a facilitation or inhibition effect cannot be attributed solely to the retrieval stage or the comparison stage.)

To examine possible strategies for processing the three-item display, the first analysis tested for an effect of position of the nonword in the display when it was with two unassociated words. A highly significant linear effect of position was found, indicating that in the majority of cases, S processed the material starting with the top item, and then proceeded serially down through the display. This appeared to continue until encountering a nonword terminated the search with a negative response, or until all three items were judged words, producing a positive response.

For positive responses including two associated words, there was a significant association effect regardless of the position of the two associated words in the display. This argues against the Location-Shifting Model.

For negative responses, the association facilitation effect occurs only if the two associated words are in positions one and two of the display. In addition to supporting the serial processing hypothesis, this provides evidence against the Semantic-Comparison Model, which predicts that the negative response for two associated words and a nonword should take longer than the response to two unassociated words with a nonword. (It could, however, be argued that the negative response on which Schaeffer and Wallace (1970) base their prediction of inhibition of associated words is not the same as the negative response used here.)

Since the interpretation of results is based to a great extent on the serial processing assumption, the finding that two nonwords followed by a word was classified "no" faster than a nonword followed by two words should be noted. Strict serial processing cannot account for this result, which suggests some kind of parallel, or overlapping processing.

The results in general do not support a Location-Shifting or Semantic-Comparison Model, and do support a Spreading-Excitation Model of associative facilitation. Also, there is evidence that the facilitation effect lasts at least 200-400 msec.

Meyer, Schvaneveldt, and Ruddy (1972) have repeated the basic experiments discussed above (two-item and three-item displays) with one important methodological change. Rather than a simultaneous display of letter strings, they are presented successively, contingent upon word/nonword classification of each item in sequence. Thus serial processing

can be assured. The basic results of associative facilitation were replicated, and it was also shown that the facilitation effect decreases by about one-half if a delay of four seconds is imposed between presentation of associated words. This also supports a Spreading-Excitation Model, with a provision for decay of excitation as a function of time. It was also shown that a visual mask (which degrades a stimulus) slows recognition less when the stimulus word is associated with a preceding word. This suggests the possibility of associative facilitation of some kind during the encoding stage (Sternberg, 1967).

Content-Addressable Models

An assumption of content-addressable models (see Norman, 1969) is that the stimulus item is the "address" of the material in memory. The studies reviewed have shown that information abstracted from the external form of the item interacts with stored information and various memory characteristics to provide the address, and suggests that the address should be considered the coordinates of a particular subset of the internal lexicon. Morton (1965) has attempted to develop a content-addressable memory system capable of accepting enough input information to find the exact "address" of a word in memory,

The theory is based on a hypothetical memory unit called a logogen (from logos--word, and genus--birth). According to Morton,

The logogen is a device which accepts information from the sensory analysis mechanisms concerning the properties of linguistic stimuli and from context-producing mechanisms. When the logogen has accumulated more than a certain amount of information, a response . . . is made available. /In this context, an available response is a word which is implicitly available for verbalization./ Each logogen is in effect defined by the information it can accept and by the response it makes available. Relevant information can be described as the set of attributes . . . semantic, visual, and acoustic sets . . .

incoming information has only a numerical effect upon any logogen which merely counts the number of members of its defining set which occur, without regard to their origin. When the count rises above a threshold value, the corresponding response is made available (pp.165-166).

The most important aspect of this model is the fact that input is accepted directly by the "storage location," the logogen. In this sense, the model is a direct-access, content-addressable model. But it is important to note the assumption that input from one word may raise the mean count in several logogens due to the fact that some words have common features. Another important assumption is that over a long time interval, the logogen counts decay to some minimum "mean" value analogous to a baseline value, unless new input is accepted by the logogen.

The important features of the model involve factors that change the threshold and/or the mean count of the logogen. Context (c) is assumed to raise the mean count of a logogen. This has the net effect of moving the total count of the logogen toward the critical value. A logogen also accepts stimulus (s) input from the set of attributes associated with a particular word. Input of these attributes raises the count of the logogens sensitive to the particular attributes. However, the stimulus effect is momentary. This assumption is required since many words have similar construction. Stimulus properties of these items would raise the count in several logogens simultaneously. If the effect of stimulus did not disappear rapidly (Morton assumes decay is complete in less than one second) the mean count of many logogens would remain high and inappropriate responses would result from input of attributes which by chance would cause the logogen to fire. Thus Morton assumes that the effect of stimulus is transient, and the effect of context is self-sustaining. The combined effect of

stimulus and context is assumed to be momentarily additive such that the count in the appropriate logogen is raised by $(c + s)$.

The effect of word frequency on a logogen is quite complex. This effect is on the threshold of a logogen, rather than mean count. Both the effect of word frequency (frequency of the type estimated by the Thorndike-Lorge 1944 count) and the momentary effect of repetition of a word are due to the same hypothesized property of a logogen. When a word is presented, the threshold of the appropriate logogen is assumed to be lowered drastically--thus the same word repeated immediately would require much less input to fire the logogen. However, within a matter of minutes after firing, threshold moves back up to a point just below the original value. The hypothesized reason that high frequency words (Thorndike-Lorge frequency) have lower thresholds relative to low frequency words results from the cumulative effect of each single presentation of a word--each time a word fires a logogen, the net effect (on threshold) is a minutely lowered threshold. Therefore more presentations of a word result in higher frequency of experience, and lower threshold according to the logogen model.

The logogen model takes into account the effect of context, specific stimuli, frequency, and repetition of the same item, and predicts their effect on the logogen. One advantage of this model is the ease with which it can be modified to account for the effects of additional variables. If a potential variable can be characterized as a set of attributes (e.g., visual patterns, acoustic patterns) provision for the variable can be built into the model. A logical extension of this notion also leads to the conclusion that the degree of precision with which these attributes can be identified is directly related to the

accuracy of predicting the response which becomes available from the logogen system. Similarly, the model may be extended such that the degree of attribute definition might also determine the value of other predicted response variables, e.g., latency of the response which becomes available. However, the logogen model does not provide for correct classification of nonwords as it was presented. And although addition of a complex decision system for producing nonword responses is possible, certain empirical results such as effects of consonant cluster frequency would be quite difficult for the logogen model to explain.

A Random Retrieval Model

An example of an entirely different model of memory organization has been proposed by Landauer (1972). This model suggests that memory is completely randomly organized, that is, each experience is laid down in memory according to the momentary location of a hypothesized "pointer" in memory. This pointer moves randomly anywhere and everywhere through memory. Retrieval is also hypothesized to be completely random. When search for a given word commences, the area within a fixed radius of the pointer is searched for the word. If the word is not found before the pointer moves on, search will continue in the next area, and so on. Although somewhat unorthodox, this model can explain many memory phenomena. For example, consider the word frequency effect. The model says that since each experience is "recorded" at a location determined by the movement of a random pointer, the more often a word is experienced the greater the probability that a representation of the given word is recorded at or near the pointer at any given time; words experienced on very few occasions would have a low probability of being located near the pointer at any given moment.

Although this model can account for many aspects of memory and retrieval, it would appear to require some additional characteristics to be able to account for correct classification of nonwords, and for differences in classification of nonwords which are a function of frequency of subgroups of letters or the legality of nonword construction.

Immediate Antecedents of the Study

Forbach et al. (1973) looked at the effects of repetition of letter strings on word-nonword decision latency. They found that classification latency for words decreased as a function of repetition. There was, however, no corresponding decrease in decision latency for nonwords. This was interpreted as an indication that the activation of a word in memory "primes" that word temporarily, such that a subsequently attempted retrieval is facilitated. The lack of a repetition effect for nonwords is congruent with the hypothesized method of nonword classification, assumed to be the exhaustive search of a subset of memory which would contain the item if it were a word. Since there is nothing stored in memory for nonwords, there should be no "priming."

Schvaneveldt and Meyer (1971), and Meyer et al. (1972) found evidence for a kind of semantic priming of associated words. Both of these findings suggest that storage and retrieval of words is closely linked to the relationship of a word and its meaning. Since nonwords generally have no meaning, they would not be expected to be stored in memory. However, if Ss were asked to learn nonsense words so they could respond positively that they were in memory, would the nonwords be integrated into the word storage network? Or would nonsense words be stored

separately in some sort of scratch file since they have no meaning? If a separate storage area is developed for nonwords, are its parameters the same as memory for words? For example, is the latency of response for indexing nonword memory the same as for word memory? Since nonwords are typically not useful past the context of an experiment in which they are learned, will nonword memory be "erased" shortly after its usefulness is past? The present study will attempt to answer these questions by comparing latencies of words and nonsense words in a memory referencing task.

Rationale

It was felt that a reasonable test of the hypothesis concerning different storage methods for words and nonwords could be made if response latency for these types of material was compared at two different times after learning. However, an important methodological problem involves the comparison of positive and negative responses to items--negative responses require more time than do positive responses. A lexical decision task is unsatisfactory then, because the critical comparison would be between positive and negative responses. Therefore, it was decided that the S would respond positively or negatively to the question "Is this letter string in your memory?" Then, after the nonwords were learned, they would be responded to positively, as would the words.

In order to make it easier to equate amount of learning for the two types of materials, both types should be unknown before the experiment. If very rare words were selected as stimuli, learning trials for the experimental group (words and meaning) could be equated with learning

trials for the control group (the same words as the experimental group, but without definitions; thus they are functionally nonsense words). Before learning trials, it would also be necessary to have the Ss classify the items as in memory or not in memory, to be certain that they did not know them as words.

If differences in memory storage exist for words as compared to nonsense words, it was felt that they would be most likely to be manifested as an interaction of type of material learned with amount of time after learning. For example, one possibility is that nonsense word memory is temporary, such that latency of response may be relatively fast immediately after learning but become slower with the passage of time and loss of material from memory. If word memory stays the same or improves after learning, an interaction would result.

Another possibility is that nonsense words are stored separately from words but remain in memory relatively permanently, such that there is no difference in classification latency after 48 hours. In conjunction with this pattern for nonwords might occur a pattern for words which decreases over the two sessions. Such a decrease could occur if the words undergo some process of consolidation between the two sessions. This overall pattern of words and nonwords would also produce an interaction.

A third possibility is that classification latency for words remains the same across the two sessions, but response latency for nonwords decreases across the two session. This pattern might occur if the S gradually learns to first check the area in which the nonsense words are stored, since the number of items to be searched through there would be minimal; it would be likely that the S would learn to do this with practice, and thus be more likely to do so during the final session.

CHAPTER II

METHODOLOGY

Subjects

A total of 41 psychology students served as volunteer subjects (Ss). Each S received a specified number of bonus points to be added to his final grade as inducement to participate. The data from one S were discarded because he was not able to attend the third session. Failure to meet the learning criterion resulted in the rejection of two other Ss. The results are therefore based on 38 Ss, nineteen each in the experimental and control groups. With the exception of replacements, Ss were assigned randomly to one of the two groups. By chance there were six males and thirteen females in each group.

Apparatus

The core of the apparatus was an eight channel Lafayette timer (Bank Timer 1431A) which controlled the timing sequence and the other equipment. Stimulus materials were presented by a Kodak Carousel projector with a five inch f3.5 lens which was equipped with a Lafayette I-24 solenoid operated shutter (power supply--Lafayette Tachistoscope VS1-E). Timing of latencies was done by a Lafayette digital Clock/Counter (Model 54417) which measured latencies to the nearest millisecond. The timing equipment was in a room adjoining the S's room, the dividing wall fitted with a one-way mirror of the

dimensions 50 x 70 cm., so that S could be observed while doing the task.

The S's room was approximately 2 x 3 m., and was painted black to minimize ambient light reflection. The S was seated at a small table at a distance of approximately 50 cm. from a 18 x 13.5 cm. Plexiglas screen onto which the materials were backprojected to produce a visual angle of approximately 4°. In his nonpreferred hand, S held a small thumb switch which initiated each trial when he was ready. The "in memory/not in memory" responses were given via a lightly sprung toggle-type switch (normally open) in a circuit with a latching relay which controlled the recording of the latencies. The switches (for right- or left-handed Ss) were mounted into the table top in such a way that S's forearm and elbow rested comfortably on the table.

A photoelectric cell responded to a light/dark spot projected out of S's view, thereby enabling a logic circuit which decoded S's response as "correct" or "incorrect." Decoding was via a red or green light in the equipment room. For the selection session, "correct" responses were defined as words which S classified as "in memory" and nonwords which S classified as "not in memory." Incorrect responses were defined as words which S classified as "not in memory" and nonwords which S classified as "in memory." For the control group in the learning and test sessions, "correct" responses included "in memory" responses to the items learned as nonwords, and "incorrect" responses included "not in memory" responses to these items. The S did not receive feedback during the experimental sessions.

Materials

Two types of stimuli were used, ~~consonant-consonant-vowel-consonant-~~consonant words, and nonwords of the same ~~consonant-vowel~~ configuration. Three categories of words were constructed based on word frequency information from the Thorndike-Lorge (1944) G count (frequency in a one million word sample). In the group of high frequency words (HF words) there were 24 words with frequency between 31 and AA. Using A=50 and AA=100, the mean of HF words was 70.3 and the median was 50.0. The 24 low frequency words (LF words) included words ranging in frequency from 1 to 17, with a mean of 7.6 and median of 4.5. Words in the final category of words, rare words (R words), were chosen from The Random House Dictionary of the English Language, The Unabridged Edition (1967) such that they were represented in neither the Thorndike-Lorge G count, nor the four million word sample. After initial selection of R words, the list was reduced to 30 by deleting those with obvious common word associates. These 30 items were given to 96 Introductory Psychology students who rated them for "meaningfulness." (See Appendix A for the rating instructions.) The final list of R words was made up of the 24 words rated lowest in "meaningfulness."

The 72 nonwords (CCVCCs) used were constructed by concatenating a consonant pair which occurs as an initial pair in English, a vowel, and a consonant pair which occurs as a terminal pair in English. The CCVCCs were then randomly assigned to one of three groups of 24 items each. It was necessary to include 48 words in the selection materials to fix the probability of an item requiring an "in memory" response at 0.50. The choice of sets of HF words and LF words was based on frequency differences found in earlier lexical decision studies. Also, the

present study provides an opportunity to compare the mean latency of "in memory" responses to the mean latency of "in vocabulary" responses found in several previous studies (Stanners et al. 1971; Stanners and Forbach, 1973; and Forbach et al. 1973). A list of all the stimulus items is available in Appendix B.

All the materials were typed in upper case with an IBM Sign typewriter, reproduced onto transparencies by the diazochrome method, and mounted in 35 mm slide holders.

Procedure

The experiment was divided into three sessions--selection, learning and test. The amount of time between the selection and learning sessions varied from 24 hours to one week. The amount of time between the learning and test sessions was always between 47 and 49 hours. The Ss were instructed with regard to this critical time period to insure their prompt arrival when scheduled.

Selection

After verifying appointments for the last two sessions, S was seated at a small table and listened to tape recorded instructions (see Appendix C). The experimenter (E) pointed out the necessary switches and lights and clarified the instructions after the tape was finished if the S so requested.

A trial was begun by S pressing a thumb switch held in his nonpreferred hand. The S had been instructed that before he pressed the switch he should attend closely to the screen and hold between thumb and forefinger of his preferred hand a lightly sprung toggle-type switch.

Following activation of the thumb switch by one second, one item was presented on the screen. The S was instructed to indicate by the direction (left or right) of the switch movement whether the item as a complete unit was "in his memory," or "not in his memory." The direction of the movement was indicated on a sign next to the switch and was held constant throughout the experiment for a given S. Each item remained on the screen until S made his response. The Clock/Counter started with the presentation of the item and stopped with S's response. Both speed and accuracy were stressed by the instructions. The offset of a small lamp below the screen three seconds after the S's response signalled that he could begin a new trial whenever he was ready.

The classification trials in all three sessions were preceded by 40 practice trials with material similar to the experimental materials. The first session lasted approximately thirty minutes.

The primary purpose of the selection trials was to choose those R words which Ss indicated were not in memory. From these, items were selected to be used in the learning session. Since not all Ss responded negatively to the same R words, each experimental S was paired with a control S with regard to the specific items chosen for learning trials. From the "not in memory" R words classed by each pair of Ss, twelve were randomly chosen for the learning session. The probability that any two pairs of Ss learned the same items was very low, but all Ss learned twelve items sampled from the same population of 24 R words.

Learning

Both groups of Ss were first read a set of instructions specific to their task (see Appendix D). For the experimental group, the words to

be learned were typed in upper case in the upper left corner of a 3 x 5 index card. A short definition for each word was typed below the word (see Appendix E for sample card layout and definitions used). The cards for the control group had only a word typed on them. Each word was assumed to be a nonsense word by the control Ss. No guide to pronunciation was given to either group.

The learning trials for both groups were paced by the audible click of a relay closure in a Hunter timer which was wired to recycle every five seconds. The deck of cards with the twelve items to be learned was shuffled before each learning trial. Each trial consisted of going through the complete deck once, with five seconds for study of each card followed by five seconds of rehearsal for each card. Each S recieved a total of five learning trials. The intertrial interval consisted of approximately one minute of conversation between E and S.

After the learning trials, each S was shown 24 cards with letter strings typed in upper case on them. Twelve of the items were the R words they had learned. The other twelve items were nonsense words which they had not previously seen. They were requested to verbally identify each as "in their memory" or not. All Ss correctly recognized at least ten study items and made no more than two false recognitions. This insured that the two groups recognized the items equally well.

The final portion of the second session consisted of 48 classification trials, including the twelve learned R words and twelve HF words as positive responses, and 24 nonwords (negative responses) which they had not previously seen. The session lasted approximately 45 minutes.

Test

The final session also included 48 classification trials. The items requiring a positive response were the same as in the preceding session, and the negative response items were 24 new nonwords. For Ss in the experimental group, a recall task followed in which they attempted to recall the definitions of the twelve R words which they had learned. A minimum retention of 75% was set to be able to assume that the learning trials in the second session had been effective. Two Ss failed to meet this criterion and were replaced. The final session lasted approximately ten minutes.

Experimental Design

The major variable, type of learning (as words, or as nonsense words) was manipulated between groups of subjects. For the selection session, the factor Category, either HF words, LF words, R words, or CCVCCs was manipulated within subjects. The number of the test session (1, 2, or 3) was a within-subjects variable. Direction of switch movement was balanced between Ss. Stimulus materials were randomly ordered for presentation to each S in the selection session. Since an experimental and a control S were paired after the selection trials, each pair of Ss received a different random order of the stimulus materials for the classification trials during both the learning and test sessions.

Scoring of Data

An individual score in the majority of the data analyses was the antilogarithm of the mean of the log latencies for a given S in a given

subcondition of the experiment (subcondition transformation). Individual latencies above two seconds were considered to be indicative of a breakdown in the decision task, and were not included in the transformed scores. The purpose of the transformation was to adjust for the skewed distribution typical of latency scores. Only "correct" responses were used in the transformed scores. For the selection session, no score was based on less than 12 of 24 possible latencies and over 99% were based on 15 or more latencies. For the learning and test sessions, no score was based on less than six of 12 possible latencies, and over 97% were based on eight or more latencies. After the transformations, the data for each S was a single score for each category of material used in each of the three experimental sessions. Subsequent analyses used these transformed scores.

The data for R words in the final two sessions were also scored by item to provide an estimate of the average latency of a particular item (item transformation). This was done by tabulating the response latency of each S for each item. Within each group, there was a maximum of 19 latencies for any given item, less if any Ss had made errors. The item transformed scores were then calculated by finding the antilogarithm of the mean of the log latencies of all Ss in a group which had learned a given R word.

Unless specified, all analyses of the transformed data discussed in this paper involved the subcondition transformed scores, since on the average each score is based on more raw latencies than the item transformed scores,

CHAPTER III

RESULTS

Selection Session

The first session for all Ss was primarily concerned with selection of R words which were not in S's memory. However, the data were analysed for frequency differences and differences between words and nonwords.

Category means, category mean error latency, and the number of errors for each category are presented in Table 1. The first analysis of variance (AOV) involved the factors Groups (G) and Category (C). The main effect of G was not significant, $F(1,36)=0.01$, $p > .25$, nor was the C x G interaction, $F(3,108)=0.11$, $p > .25$. The main effect of C was significant, $F(3,108)=45.91$, $p < .001$.

The next two AOVs examined each group separately. There was a significant effect of C in each group, $F(3,54)=22.47$, $p < .001$, and $F(3,54)=23.51$, $p < .001$ for the experimental and control groups, respectively. Consequently, the Newman-Keuls procedure was employed to check for differences among category means in each group. In each group, there was a significant difference between all means, with the exception of the difference between CCVCCs and R words. This difference was significant in neither group. Table 2 summarizes the results of the tests and indicates significance level for appropriate comparisons. Since there was no overall significant difference between groups, and

TABLE I
MEAN LATENCY, MEAN ERROR LATENCY, AND NUMBER OF ERRORS
IN EACH CATEGORY FOR THE SELECTION SESSION
CLASSIFICATION TRIALS

Category	Correct Latencies		Error Latencies		Total Errors	
	<u>Exp</u>	<u>Con</u>	<u>Exp</u>	<u>Con</u>	<u>Exp</u>	<u>Con</u>
HF Words	710	722	925	655	20	11
LF Words	787	795	836	834	55	62
CCVCCs	841	866	901	814	25	28
R Words	878	893	742	786	22	40

TABLE II
SUMMARY OF SIGNIFICANT DIFFERENCES IN CATEGORY
MEANS, AS TESTED BY THE NEWMAN-KEULS TEST

Categories	Experimental				Control			
	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	<u>(4)</u>	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	<u>(4)</u>
HF Words (1)	-----	77**	131**	168**	-----	73**	144**	171**
LF Words (2)		-----	54*	91**		-----	71**	98**
CCVCCs (3)			-----	37			-----	27
R Words (4)				-----				-----

* $p < .05$ ** $p < .01$

no interaction of groups with categories, no comparisons of means were made between groups.

The error rates for the two groups were very close, 6.7% for the experimental group and 7.7% for the control group.

The mean error latencies for the word items in each group were slower than the corresponding correct latencies, with the exception of the error latency for HF words from the control group. Conversely, mean error latencies for the nonwords in each group were faster than the corresponding correct latencies, excepting the CCVCC error latency from the experimental group.

Learning and Test Sessions

A summary of the descriptive statistics for the final two sessions of the experiment is presented in Table 3. The data of primary interest are the response latencies for R words. The predicted interaction would be expected to occur if the difference in mean latency between the experimental and control groups for the learning session classification trials (a difference of 22 msec) was statistically different from the corresponding difference in group means for the test session trials (a difference of 88 msec). The mean latency of response to R words for both groups in the final two sessions is presented in Figure 1. A two-factor AOV with the factors Group (G) and Test Session (T) was employed to evaluate the reliability of the difference between 22 msec and 88 msec. Neither the main effect of G nor T was significant, $F(1,18)=0.72$, $p > .25$, and $F(1,18)=3.10$, $p < .10$ respectively. The G x T interaction was also nonsignificant, $F(1,18)=2.08$, $p < .20$.

The scores obtained from the item transformation of R words were

TABLE III

MEAN LATENCY, MEAN ERROR LATENCY, AND NUMBER OF ERRORS
IN EACH CATEGORY FOR THE LEARNING AND TEST
SESSION CLASSIFICATION TRIALS

Session	Category	Correct Latencies		Error Latencies		Total Errors	
		<u>Exp</u>	<u>Con</u>	<u>Exp</u>	<u>Con</u>	<u>Exp</u>	<u>Con</u>
<u>Learning</u>	HF Words	759	782	788	593	3	5
	R Words	862	840	908	893	14	5
	CCVCCs	879	906	883	751	27	19
<u>Test</u>	HF Words	694	684	867	937	3	4
	R Words	855	767	977	938	12	7
	CCVCCs	853	886	919	824	20	21

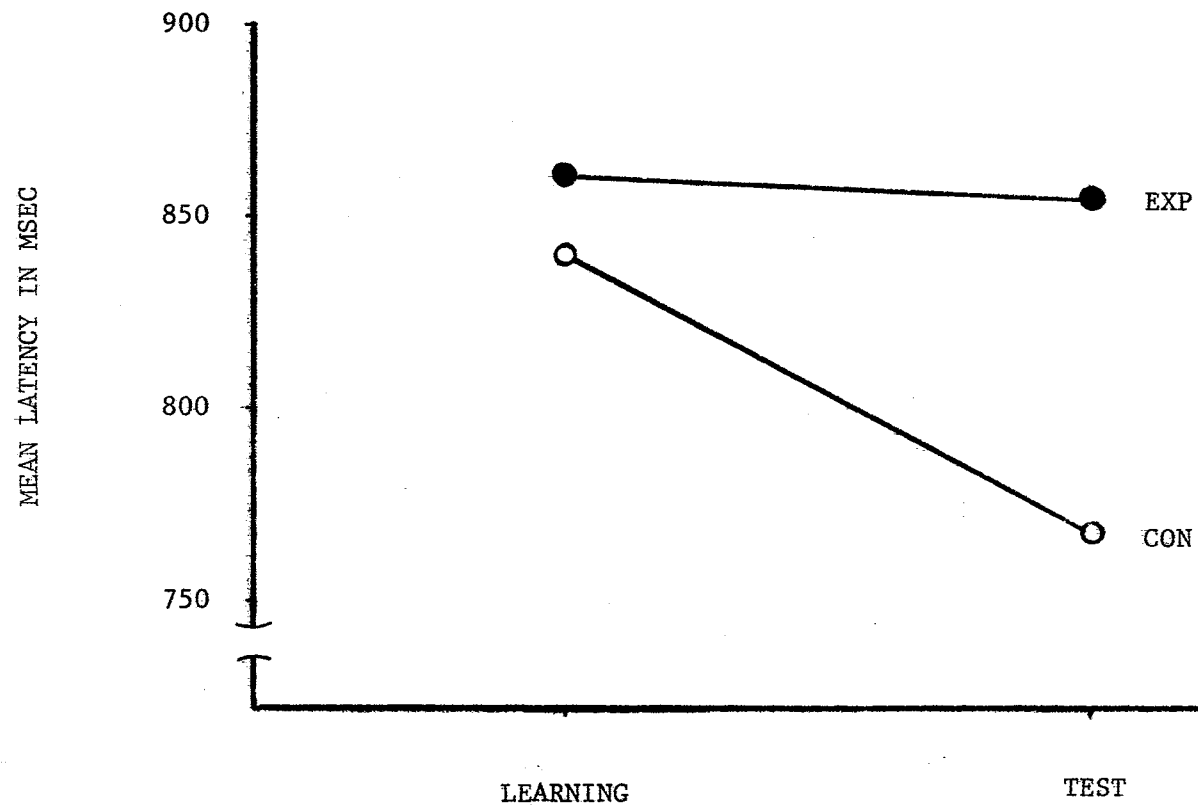


Figure 1. Mean Latency of Response for the Experimental and Control Groups Immediately After Learning R Words (Learning), and Then 48 Hours Later (Test)

also used for an analysis of the predicted interaction effect. The difference between the mean of the experimental and the control group for the learning session classification trials was 31 msec. The corresponding difference for the test session trials was 42 msec. The reliability of the difference between 31 msec and 42 msec was determined by examination of the $G \times T$ interaction term in another two-factor AOV. The interaction was not significant, $F(1,23)=0.14$, $p > .25$, nor was the main effect of G , $F(1,23)=1.97$, $p > .25$. However the main effect of T was significant, $F(1,23)=8.53$, $p < .01$. The analysis of the subcondition transformed scores had indicated that the main effect of T across sessions two and three only tended toward significance, $p < .10$ (above). This slight inconsistency was checked by examination of the homogeneity of variance assumptions for the subcondition transformed data. The variances of both groups were compared for the second and third sessions. The test indicated a significant departure from homogeneity, $F_{\max}(4,18)=3.30$, $p < .05$. Therefore, the experimental and control groups were analysed separately to check for an effect of T by comparing R word scores on the learning session trials with scores on the test session trials. For the experimental group, no improvement was made over the two sessions, $F(1,18)=0.06$, $p > .25$. However, the control group in the third session showed a significant decrease in classification latency from that in the second session trials, $F(1,18)=4.81$, $p < .05$.

A final two-factor AOV was employed to evaluate the effect of repeating in the test session trials, the twelve HF words used in the learning session classification trials. There was no main effect of G , $F(1,18)=0.03$, $p > .25$, nor a $G \times T$ interaction, $F(1,18)=1.47$, $p > .25$. There was, however, a strong effect of T , $F(1,18)=36.46$, $p < .001$,

indicating a significant decrease for both groups on the final session.

The error rates for both groups in each session were comparable. For the learning session, they were 5.1% and 3.2% for the experimental and control groups, respectively. The corresponding error rates for the test session were 3.9% and 3.5%.

Examination of the error latencies revealed that the experimental group CCVCC error latencies for both the learning and test sessions were slower than the correct latencies, and that the control group HF word latencies for the learning session trials were faster than the correct latencies. This is opposite the typical pattern reflected by the remaining error latencies, namely that word item error latencies are slower than correct word latencies, and nonword item error latencies are faster than correct nonword latencies.

CHAPTER IV

DISCUSSION

Selection Session

Examination of the results for the first session reveals two major points of interest. First, the relationships among the means of HF words, LF words, and nonwords correspond to that found by Stanners et al. (1971), Stanners and Forbach (1973), Forbach et al. (1973), and Rubenstein et al. (1970, 1971a, 1971b). Also, the expected pattern of significant differences between means which had been found to be reliably different in some of these earlier studies was found, namely that CCVCC HF and LF word means differ, and each differs from CCVCC nonsense word means. These findings lend support to search models which have been proposed in these previous studies.

Second, the amount of time required for an "in your vocabulary/not in your vocabulary" (lexical) decision task can be compared with that for the "in memory/not in memory" decision task. Stanners and Forbach (1973), and an unpublished replication of their study provides an estimate of lexical decision time which is based on data from 80 Ss. For HF words the latency is 614 msec, for LF words the average classification latency is 690 msec, and for CCVCC nonwords the average latency is 732 msec. When the task involves classification of items as "in memory" or "not in memory," as in the present study, the average classification time (based on 40 Ss is over 100 msec higher in all categories.

For HF words the average latency is 716 msec, for LF words the mean is 791 msec, and for CCVCCs the mean latency is 871 msec. This consistent difference suggests that possibly one or more stages are added to the search process, to check the area in which nonwords would be located. This interpretation is compatible with the multi-stage search models which have been proposed previously. The Ss in the present study were instructed that humans may temporarily store nonsense words in memory (e.g., if the S had just seen them or heard them previously). The instructions in the lexical decision studies made no mention of the possibility that nonwords could be stored in memory, and in fact the Ss were instructed to respond positively only to words in their vocabulary (see Appendix F for these instructions). Since the only obvious differences in the two types of classification tasks is associated with the difference in instructions, some kind of "extended search" hypothesis may be tenable. This hypothesis could be tested by having Ss respond to items as "in my vocabulary" or "not in my vocabulary" for one-half of an experiment, and then respond "in my memory" or "not in my memory" for the remainder of the experiment. The required change in response would make it necessary to attend to nonwords seen previously, and to respond positively to them. (This also assumes proper control of the order in which the two methods of responding are required.) An attempted ad hoc statistical evaluation of the apparent differences associated with the two types of classification tasks would probably be of little value because of the lack of the proper design. However, the differences in the means of the present study from those of lexical decision studies suggest another potential method of demonstrating that memory for words and nonwords may be functionally different.

The pattern of error rates and error latencies found in this study had only minor deviations (see Results chapter) from that found in lexical decision studies. Since there is no evidence to the contrary, the inconsistencies in the pattern of error latencies can probably be attributed to the low number of raw scores used to estimate the error latencies.

Learning and Test Sessions

The primary goal of the experiment was to investigate the hypothesis that words and nonwords are stored in separate memory locations, each having different parameters of storage and retrieval. Partial support was given, since the trend of the data was toward the expected interaction. However, the effect was not statistically reliable ($p < .20$). Analysis of the data from the learning and test sessions within each group separately indicated that the experimental group showed no change in classification latency for R words over the final two sessions. However, a similar comparison for the control group showed a significant drop in classification latency during the test session trials, which were 48 hours after the learning classification trials. This pattern might be expected to produce a significant interaction term, and the fact that it did not suggests that all the assumptions for AOV were not met. Accordingly, the error mean square associated with each group in both the learning and test session classification trials was examined. Immediately after the learning trials, the variances for the experimental and control groups were approximately equal, 11,547.04 (s.e. of \bar{X} =107.4) and 12,747.61 (s.e. of \bar{X} =112.9) respectively. The test session variance for the experimental group increased by about one-half to 18,679.79

(s.e. of \bar{X} = 136.7). However, the test session variance for the control group decreased by about one-half to 5,661.93 (s.e. of \bar{X} = 75.3). Thus while the learning session variances appeared homogeneous, the test session variance of the experimental group was significantly larger than the variance of the control group (see Results chapter). One possible interpretation of this result is that there was much more error variance associated with the experimental group because learning of the words was not complete. During the design of the present experiment, it was felt that learning words and their definitions would be easier than learning nonsense words since the words have an existing framework into which they can be integrated. Since the actual amount of information to be learned by the experimental group was larger than the control group, but the amount of time allowed for consolidation of the material was equal for both groups, it is not unreasonable to expect that there could be differences in the amount of material that reached relatively permanent memory storage. Apparently, the number of learning trials for the experimental group was insufficient to integrate them into memory properly. This suggests that the learning trials should be extended, or other methods of learning might be used. For example, the Ss might be asked to generate sentences in which the words are used. Or the study trials could include reading paragraphs or sentences which use the words. Another possible method (similar to one way we learn words) would involve listening to a tape recorded discourse involving the words. Additional integrative facilitation should occur from context, pitch changes, phrasing, emotional expression, etc. These kinds of training trials would probably increase the amount of material that reaches a more permanent memory store. Apparently, learning a new vocabulary

word to the extent that it becomes a relatively permanent part of memory involves much more than 50 seconds of study of a short definition.

The fact that response latency for the control group actually decreased significantly for the final session is somewhat perplexing. One possible explanation involves that alluded to in the earlier discussion of possible outcomes of the experiment, namely a change in search strategy by the control Ss. If there are in fact separate storage locations for nonwords, as the S has more and more practice on the task he may decide to check the nonword storage area first on all trials. If he gradually learns to check the nonword area first, the lowest latencies would be expected on the trials in the last session. However, if this were the case, the control group latencies for words should be slower than the experimental group word latencies. Since there were no significant differences in word latencies between the two groups, this interpretation is probably untenable.

There does appear to be an appropriate alternative explanation of the decrease in nonword latency found in the final session. The effect appears to be related to an extensively reported phenomenon in serial learning literature called reminiscence (Buxton, 1943). Although there is no complete agreement as to the conditions required to produce the effect, there is widespread evidence supporting reminiscence, including a recent paper by Scheirer and Voss (1969). Reminiscence is generally defined as an improved performance on retest trials given after a rest period, without benefit of any practice or additional learning trials subsequent to the first test trials. In the present experiment, Ss classified R words immediately after learning trials. They had no further opportunity to study or practice them (unless they rehearsed

them on their own in the absense of any stimuli). Performance by Ss in the control group improved significantly on the trials in the last session, suggesting a reminiscence effect. It is interesting to note that one of the original studies which found the effect (Hovland, 1938) used nonsense syllables as stimuli. Subsequent attempts to demonstrate a reminiscence effect for words included an unsuccessful attempt by Melton and Stone (1942). In the present study, the effect occurred for nonsense materials (control group), but not for words (experimental group).

The slight inconsistency in the pattern of error latencies (see Results chapter) found in the present study for the final two sessions was not considered indicative of some difference in the performance of the Ss in the present study from Ss in lexical decision studies. It is probably due to the low number of raw scores which were used to develop mean error latencies.

CHAPTER V

SUMMARY

The present study fell somewhat short of providing statistically reliable evidence to support the hypothesis that words and nonwords are stored differently in relatively permanent memory. The trend of the data was in the predicted direction, thus lending some encouragement to the hypothesis. Indirect support for the hypothesis was obtained by comparing overall category means in the present study with those in lexical decision studies. The means in the present study are consistently higher, suggesting possibly that search of some additional storage locations (the hypothesized store for nonwords) is required in the present study.

The pattern of means found in the present study is consistent with those in earlier studies, thus supporting the general search models proposed in several lexical decision studies.

Finally, although the apparent reminiscence effect was not anticipated, it is not unreasonable given the similarities between the learning technique used in the present study, and those used in typical serial learning studies. The finding that no reminiscence effect occurred for words could also be interpreted as supportive evidence that different types of memory storage are employed for words and nonwords.

In conclusion, the present study appears to provide enough

suggestive evidence consistent with the proposed hypothesis to warrant further investigation. However, this investigation should proceed after certain methodological changes, primarily, the type and number of learning trials employed to put new words into relatively permanent memory.

A SELECTED BIBLIOGRAPHY

- Buxton, C. E. The status of research in reminiscence. Psychological Bulletin, 1943, 40, 313-340.
- Cohen, B. H., Bousfield, W. A., & Witmarsh, G. A. Cultural norms for verbal items in 43 categories. Technical Report No. 22, 1957, University of Connecticut, Contract Nonr 631(00), Office of Naval Research.
- Collins, A. M., & Quillian, M. R. Facilitating retrieval from semantic memory: The effect of repeating part of an inference, Acta Psychologica, 33, Attention and Performance III (A. F. Sanders, ed.) 1970, 304-314.
- Forbach, G. B., Stanners, R. F., & Hochhaus, L. Repetition and practice effects in a lexical decision task. Unpublished manuscript, Oklahoma State University, Stillwater, Oklahoma, 1973.
- Hovland, C. I. Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and by distributed practice. Journal of Experimental Psychology, 1938, 22, 201-224.
- Howes, D. On the relationship between the probability of a word as an association and in general linguistic usage. Journal of Abnormal Social Psychology, 1957, 54, 75-85.
- Johnson, D. M. Word-association and word-frequency. American Journal of Psychology, 1956, 69, 125-126.
- Kollasch, S. F., & Kausler, D. H. Recognition learning of homophone. Journal of Experimental Psychology, 1972, 92, 432-434.
- Landauer, T. K. Memory without organization: Some properties of a model with random storage and undirected retrieval. Paper presented at the meeting of the Psychonomic Society, St. Louis, Mo., November, 1972.
- Loftus, E., Freedman, J. L., & Loftus, G. R. Retrieval of words from subordinate and superordinate categories in semantic hierarchies. Psychonomic Science, 1970, 21, 235-236.
- Mayzner, M. S., & Tresselt, M. E. Tables of trigram frequency counts for various word-length and letter-position combinations. Psychonomic Monograph Supplements, 1965, 1, 33-78.

- Melton, A. W., & Stone, G. R. The retention of serial lists of adjectives over short time-intervals with varying rates of presentation. Journal of Experimental Psychology, 30, 1942, 295-310.
- Meyer, D. E., & Ellis, G. B. Parallel processes in word recognition. Paper presented at the annual meeting of the Psychonomic Society, San Antonio, Texas, November, 1970.
- Meyer, D. E., & Schvaneveldt, R. W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. Journal of Experimental Psychology, 1971, 90, 227-234.
- Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Activation of lexical memory. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, Mo., November, 1972.
- Morton, John. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.
- Norman, D. A. Memory and Attention. Wiley: New York, 1969.
- The Random House Dictionary of the English Language, the Unabridged Edition. New York: Random House, 1967.
- Rubenstein, H., Garfield, L., & Millikan, J. Homographic entries in the internal lexicon. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 487-494.
- Rubenstein, H., Lewis, S., & Rubenstein, M. Homographic entries in the internal lexicon: Effects of systematicity and relative frequency of meanings. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 57-62. (a)
- Rubenstein, H., Lewis, S., & Rubenstein, M. Evidence for phonemic recoding in visual word recognition. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 645-657. (b)
- Schaeffer, B., & Wallace, R. The comparison of word meanings. Journal of Experimental Psychology, 1970, 86, 144-152.
- Scherier, C. J., & Voss, J. F. Reminiscence in short-term memory. Journal of Experimental Psychology, 1969, 80, 262-270.
- Schvaneveldt, R., & Meyer, D. Retrieval and comparison processes in semantic memory. Paper presented at the Fourth International Symposium on Attention and Performance, Boulder, Colorado, August 16-21, 1971. In Attention and Performance IV, S. Kornblum (ed.). New York: Academic Press, in press.
- Stanners, R. F., Forbach, G. B., & Headley, D. B. Decision and search processes in word-nonword classification. Journal of Experimental Psychology, 1971, 90, 45-50.

- Stanners, R. F., & Forbach, G. B. Analysis of letter strings in word recognition. Journal of Experimental Psychology, 1973, 92, 31-55.
- Sternberg, S. Two operations in character recognition: Some evidence from reaction-time measurements. Perception and Psychophysics, 1967, 2, 43-53.
- Stroop, J. R. Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 1935, 18, 643-661.
- Thorndike, E. L., & Lorge, I. The Teacher's Word Book of 30,000 Words. Bureau of Publications Teachers College, Columbia University: New York, 1944.
- Underwood, B. J., & Schulz, R. W. Meaningfulness and Verbal Learning. Chicago: Lippincott, 1960.
- Warren, R. E. Stimulus encoding and memory. Journal of Experimental Psychology, 1972, 94, 90-100.

APPENDIX A

MEANINGFULNESS RATING INSTRUCTIONS

FOR R WORDS SELECTION

The following instructions were read before rating the items.

Please read the complete list of letter strings included before starting; this will give you a general idea about the items with which you will be working.

After reading the list, please go back through the list and rate each one individually for "meaningfulness." Use the following scale:

a	b	c	d	e
0-20	21-40	41-60	61-80	81-100

A rating of "a" would be the lowest "meaningfulness" value, and a rating of "e" would signify the highest "meaningfulness."

Your rating of "meaningfulness" will necessarily be very subjective and impressionistic. Since probably none of the letter strings are words, there is no "meaning" in the usual sense. However, each item may remind you of other items or look as if it should have a particular meaning. If you feel this way about a particular item, try to gauge the "amount" of "meaningfulness" and rate it appropriately.

Since you read the complete list of items before starting you should have a feel for the spread of "meaningfulness" represented in this sample of letter strings. Try to use the whole scale to rate the items even though these particular ones may not be as high or low as some other letter strings might be. In other words, try to consider the complete scale, and consider only these items using the full scale to rate them.

APPENDIX B

STIMULUS ITEMS USED IN EACH CLASSIFICATION SESSION

	<u>Selection Session Stimuli</u>			
R Words	BRACT	CROFT	CRUCK	FLUMP
	FREMD	FRITH	GLISK	GLOST
	GRIFF	PRINK	SHENT	SKELP
	SLOJD	SMARM	SNECK	SPALL
	STECH	STIRK	STOSS	STURT
	SWARF	SWITH	THEGN	TRANK
CCVCCs	BRESK	BRILD	CHENK	CLAFF
	CLENT	DRELN	DROFF	DWERK
	FRALT	FROCH	GRAST	KNARN
	PLAPT	PLENK	PNACK	PRAST
	SCURN	SHART	SHENG	SLENT
	TROFT	TRULL	WHEPT	WHICT
LF Words	BLAND	BLEND	BLINK	BRAWN
	CHANT	CHESS	CHOCK	CLASH
	CLICK	CRESS	FLICK	GLAND
	GNASH	GRAFT	GRAPH	GRILL
	GROSS	GRUNT	KNACK	PRONG
	SHACK	TRACT	WRIST	WHELP
HF Words	BLAST	BLESS	BLIND	BRASS
	BROWN	CHECK	CLOCK	CREPT
	CROSS	DRESS	FLASH	FLOCK
	GLASS	GRAND	GRANT	GRASS
	SHOCK	STAMP	STICK	STILL
	SWELL	THING	TRACK	TRUTH

Learning Session Stimuli

R Words

(12 of the items in the 24-item pool)

CCVCCs

CHOLD	DWALT	DWILL	FRASH
FRENT	GLEND	GNEMB	GNUSH
GRESS	GROPH	KNEPT	PHAFT
PRAMP	RHOLP	STOLK	SWESH
THAMB	TRATH	TRIMP	TWEXT
TWILB	TWING	WRICK	WRILF

HF Words

BROWN	CHECK	CROSS	DRESS
FLASH	GLASS	GRANT	GRASS
STICK	STILL	THING	TRUTH

Test Session Stimuli

R Words

(Same 12 as used in learning session trials)

CCVCCs

BRUNX	CHACK	GHAMN	GLANT
GLAPH	GROFT	KNUCK	LLAMT
MNETZ	PLEMB	PLINT	PNIGN
PSEFF	PSELM	SHILF	SHULL
STIMP	SWEMP	SWUFT	TSENC
TWEST	WHECH	WHOLP	WRUTH

HF Words

BROWN	CHECK	CROSS	DRESS
FLASH	GLASS	GRANT	GRASS
STICK	STILL	THING	TRUTH

APPENDIX C

INSTRUCTIONS TO SUBJECTS FOR SELECTION SESSION TRIALS

The following instructions were tape recorded and played to all Ss before the beginning of the experiment.

This is an experiment concerned with simple judgements about verbal materials. It is not an intelligence test of any kind and should not be interpreted as such. Also, there is no electric shock nor any other unpleasant stimulus involved. Although the task may seem to be a very simple one, our research indicates that it can provide important information concerning language behavior. If for any reason during the course of the experiment you feel that you cannot fully cooperate, please let the experimenter know.

A five-letter item will be presented on the screen in front of you [E indicates]. Your job is to decide, as quickly as possible, whether the item is or is not in your memory. If you decide the item is in your memory, move the switch in the direction indicated on the card [E indicates]. If the item is not in your memory, move the switch in the opposite direction. Make your judgement on the basis of whether the item is a complete unit in your memory without adding anything to it. On this basis, the item S-P-A-R-C would not be in most people's memory, even though it is similar to and may remind you of the word S-P-A-R-K. In the same way the item S-L-A-N-D would not be in most people's memory even though it is similar to and sounds like S-L-A-N-T. First or last names should also not be treated as independent units. Examples of names which you might recognize, but should not be treated as independent units are C-H-U-C-K and S-M-I-T-H.

Of course, an exception to the independent unit rule might occur if I now repeated the items S-P-A-R-C and S-L-A-N-D. That is, since you had just previously seen them they might still be in your memory. However, unless they were being repeated, items such as these would usually probably not be in your memory.

Slang terms may be treated as independent units. If they are in your memory, you should indicate this with your response. Examples of fairly common slang terms are S-W-E-L-L and C-H-U-M-P. If items such as these are in your memory,

then you should respond appropriately.

At the beginning of the series of trials I will sound a buzzer. You can then start a trial by pressing the thumb button, which you should hold in your nonpreferred hand. About a second after you press the thumb button the item will appear on the screen and you should respond with the switch as quickly and accurately as you can. Make sure that when you press the thumb button you are paying very close attention to the screen and that you are holding the switch between the thumb and forefinger of your preferred hand / E indicates /. If you are ready to respond when you press the thumb button your switch responses will be faster. It is very important for a successful experiment that you concentrate fully on each item, and classify it as quickly and accurately as possible.

The white light below the screen will signal when you can start another trial; the thumb button will not work until after the white light goes off. You do not have to start another trial immediately after the white light goes off. If you want to take a short break, that is OK.

I will not attempt to trick or confuse you by repeating items. In the next session you attend you will learn some items and store them in your memory. Thus I am now interested only in those items which were in your memory when you reported for the experiment.

Are there any questions?

APPENDIX D

INSTRUCTIONS TO SUBJECTS FOR

THE LEARNING TRIALS

The following instructions were read to all Ss.

You have been randomly assigned to one of two groups in this experiment. Your group differs from the other only in the way you will learn some items in the next few minutes. Thus it is very important that you try to follow your specific instructions exactly. Also, it is vital for the success of the experiment that you not discuss any aspect of today's experiment with anyone, especially if they too are participating. Since I will be conducting the experiment for several weeks, discussion with other participants may bias the results in an unknown manner. This is not an attempt to trick or confuse you in any way; I will be quite willing to explain the experiment to you in whatever detail you desire after your participation has been completed. However, please do not discuss it with anyone else until all the data has been collected.

The following instructions were read to experimental Ss only.

Among the items which you classified as not in your memory previously, there were some very rare words. What I want you to do today is learn twelve of these words and their definitions. The procedure we will use is as follows: First I will show you a card with the word and its definition for five seconds. Next I will remove the card and let you rehearse the word for five more seconds. This will be repeated for all twelve of the items. We will then take a short break, and finally go through the whole list four more times in the same manner. At this point you should know them. I will then show you the twelve words mixed in with some other items and ask you to say if each is in your memory or not. Then you will have some classification trials again to see if the status of the new words has changed. At the end of the next session you attend, I will ask you to write down the definitions of the twelve words you learned today.

The following instructions were read to control Ss only.

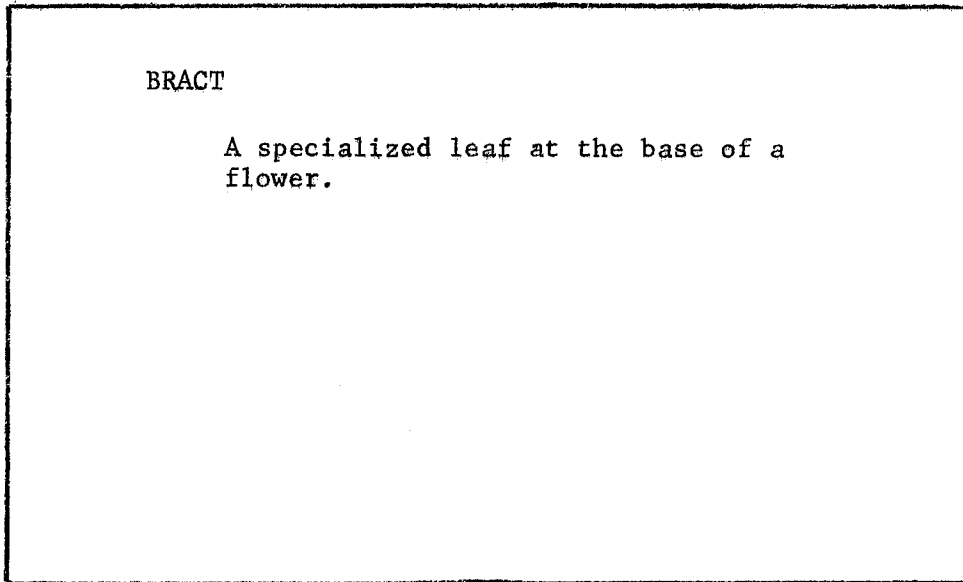
From the items which you classified as not in your

memory previously. I have chosen twelve for you to learn. The procedure we will use is as follows: First I will show you a card with the item for five seconds. Next I will remove the card and let you rehearse the item for five more seconds. This will be repeated for all twelve of the items. After a short break, we will go through the list four more times in the same manner to insure that you have learned them. Then I will show you a random ordering of 24 items, including the ones you have just learned, and ask you to classify them as in your memory or not. Finally, you will do a few more memory classification trials to see if the status of the items you studied has changed.

APPENDIX E

SAMPLE LAYOUT OF CARDS USED FOR LEARNING, AND STUDY DEFINITIONS

Cards which the Ss in the experimental group studied included a word and a short definition, such as the following sample card:



Cards which the Ss in the control group studied included only a word, such as the following sample card:

BRACK

Study Definitions for R Words

BRACK -- A specialized leaf at the base of a flower.
 CROFT -- A very small garden plot.
 CRUCK -- A naturally curved timber used as a roof support.
 FLUMP -- To flop down suddenly.
 FREMD -- Foreign or strange.
 FRITH -- A long narrow indentation of a seacoast.
 GLISK -- A small glimmer of light.
 GLOST -- A ceramic glazed finish.
 GRIFF -- A newcomer or "greenhorn."
 PRINK -- To deck or dress for show.
 SHENT -- To put to shame.
 SKELP -- A slap with the open hand.
 SLOJD -- An apprentice program in woodworking.
 SMARM -- Trite sentimentality.
 SNECK -- A door latch or lever.
 SPALL -- A chip or splinter of stone.
 STECH -- To gorge with food.
 STIRK -- A young bull or cow.
 STOSS -- Land which received the thrust of a glacier.
 STURT -- Violent quarreling.
 SWARF -- Accumulation of particles from metal grinding.
 SWITH -- To hurry or hasten immediately.
 THEGN -- A class name in early England.
 TRANK -- Leather from which a glove is cut.

APPENDIX F

INSTRUCTIONS TO SUBJECTS IN LEXICAL DECISION STUDIES

The following instructions were tape recorded and played to all Ss before beginning the experiment.

This is an experiment concerned with simple judgements about verbal materials. It is not an intelligence test of any kind and should not be interpreted as such. Also, there is no electric shock nor any other unpleasant stimulus involved. Although the task may seem to be a very simple one, our research indicates that it can provide important information concerning language behavior. If for any reason during the course of the experiment you feel that you cannot fully cooperate, please let the experimenter know.

A five-letter item will be presented on the screen in front of you / E indicates /. Your job is to decide, as quickly as possible, whether or not the item is or is not part of your vocabulary. If you decide the item is in your vocabulary, move the switch in the direction indicated on the card / E indicates /. If the item is not part of your vocabulary, move the switch in the opposite direction. Make your judgement on the basis of whether the item is a complete unit in your vocabulary without adding anything to it. On this basis, the item S-P-A-R-C would not be a member of most people's vocabulary, even though it is similar to and may remind you of the word S-P-A-R-K. In the same way the item S-L-A-N-D would not be in most people's vocabulary even though it is similar to and sounds like S-L-A-N-T. First or last names should also not be treated as independent units. Examples of names which you might recognize, but should not be treated as independent units are C-H-U-C-K and S-M-I-T-H.

Slang terms may be treated as independent units. If they are members of your vocabulary, you should indicate this with your response. Examples of fairly common slang terms are S-W-E-L-L and C-H-U-M-P. If items such as these are part of your vocabulary, then you should respond appropriately.

A complete trial sequence will proceed like this: You should hold the thumb button in your nonpreferred hand / E indicates /. Also, you should hold the switch between the thumb and forefinger of your preferred hand / E indicates /.

When the experimenter is in the next room, and ready to start, a buzzer will sound indicating that you may begin. Start each trial by pressing the thumb button. About one second after you press it, the item will appear on the screen. As quickly as possible, decide whether the item is in your vocabulary or not, and move the switch in the appropriate direction. Both speed and accuracy are important. After your response, move the switch back to the middle position. Make sure that when you press the thumb button you are paying very close attention to the screen and that you are holding the switch properly. After your response, the white light will come on for a short rest interval. You may not activate the next trial until the white light goes off. After the white light goes off, you may start another trial when you wish, making sure you are paying very close attention to the screen before you press the thumb button.

Are there any questions?

VITA

Gary Bruce Forbach

Candidate for the Degree of

Doctor of Philosophy

Thesis: FUNCTIONAL DIFFERENCES IN THE STORAGE OF WORDS AND NONWORDS
IN LONG-TERM MEMORY

Major Field: Psychology

Biographical:

Personal Data: Born September 8, 1947, the son of Mr. and Mrs.
A. L. Forbach, in Casa Grande, Arizona.

Education: Graduated from Eastwood High School, El Paso, Texas,
in May, 1965; received Bachelor of Arts degree in Psychology
from University of Denver, Denver, Colorado, in June, 1969;
received Master of Science degree from Oklahoma State Univer-
sity in Psychology, July, 1972; enrolled in doctoral degree
program, August, 1972 at Oklahoma State University, and
finished degree requirments in July, 1973.

Professional Experience: Served as Graduate Teaching Assistant,
College of Arts and Sciences, Oklahoma State University,
1969-1970; served as Graduate Research Assistant 1970-1971;
was National Science Foundation Summer Research Assistant,
1971; also, Graduate Research Assistant, August 1971 through
July 1972; served as Graduate Research Assistant August,
1972 through May, 1973; member Southwestern Psychological
Association; hired as Assistant Professor of Psychology by
Washburn University of Topeka, Kansas, to begin teaching
duties August, 1973.